(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 743 774 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
20.11.1996 Bulletin 1996/47

(51) Int. Cl.⁶: H04L 9/08, H04L 9/32

(21) Application number: 96201322.3

(22) Date of filing: 17.05.1996

(84) Designated Contracting States:
CH DE ES FR GB IT LI NL SE

(30) Priority: 18.05.1995 GB 9510035

(71) Applicant: Certicom Corp.
Mississauga, Ontario L5R 3L7 (CA)

(72) Inventors:
- Vanstone, Scott A.
  Waterloo, Ontario N2T 2H4 (CA)
- Menezes, Alfred John
  Auburn, Alabama 36830 (US)
- Qu, Minghua
  Waterloo, Ontario N2L 3E5 (CA)

(74) Representative: de Vries, Johannes Hendrik Fokke
De Vries & Metman,
Gebouw Autumn,
Overschiestraat 184 N
1062 XK Amsterdam (NL)

(54) **Strengthened public key protocol**

(57) A cryptosystem utilizes the properties of discrete logs in finite groups, either in a public key message exchange or in a key exchange and generation protocol. If the group selected has subgroups of relatively small order, the message may be exponentiated by a factor of the order of the group to place the message in a subgroup of relatively small order. To inhibit such substitution, the base or generator of the cryptosystem is chosen to be a generator of a subgroup of prime order or a subgroup of an order having a number of relatively small divisors. The message may be exponentiated to each of the relatively small divisors and the result checked for the group identity. If the group identity is found, it indicates a vulnerability to substitution and is rejected.

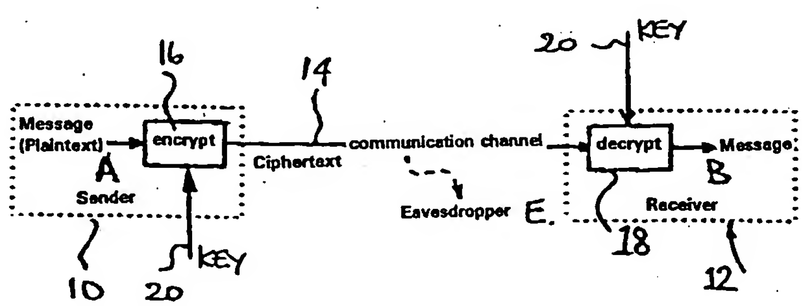

Fig. 1

EP 0 743 774 A2

Description

The present invention relates to public key cryptography.

It is well known that data can be encrypted by utilising a pair of keys, one of which is public and one of which is private. The keys are mathematically related such that data encrypted by the public key may only be decrypted by the private key. In this way, the public key of a recipient may be made available so that data intended for that recipient may be encrypted with the public key and only decrypted by the recipients private key.

One well-known and accepted public key cryptosystem is that based upon discrete logarithms in finite groups. Different finite groups may be used, for example the multiplicative group Z_p^* of integers mod p where p is a prime; the multiplicative group of an arbitrary finite field e.g. $GF2^n$ or an elliptic curve group over a finite field.

The discrete log problem used in such cryptosystems is based on the difficulty of determining the value of an integer x from the value of α^x, even where α is known. More particularly, if α is an element of G (which is considered to be written multiplicatively) and β is a second element of G, then the discrete logarithm problem in G is that of determining whether there exists an integer x such that $\beta = \alpha^x$, and if so, of determining such a value x.

The Diffie-Hellman key exchange protocol is widely accepted and there are numerous examples of implementations of the Diffie-Hellman protocol in use around the world.

The Diffie-Hellman key agreement protocol is typically stated as follows using as an example the finite group Z_p^*:

SETUP

The protocol requires a base α that generates a large number of elements of the selected group G and a pair of integers x,y that are retained confidential by respective correspondents A,B. Select a prime number p and let α be a generator of the multiplicative group Z_p^*, i.e. the group of integers modulo p.

THE PROTOCOL

1. Correspondent A generates a random integer x, computes α^x and sends this to correspondent B.
2. Correspondent B generates a random integer y, computes α^y and sends this to correspondent A.
3. A computes $(\alpha^y)^x = \alpha^{xy}$.
4. B computes $(\alpha^x)^y = \alpha^{xy}$.

A and B now share the common key α^{xy} which may be used as a secret key in a conventional cryptosystem. A similar protocol may be used in a public key system, generally referred to as an E1-Gamal protocol in which each correspondent has a secret key x and a public key α^x.

The security of these protocols seems to rest on the intractability of the discrete logarithm problem in the finite group G. It should also be noted that the protocol carries over to any finite group.

The applicants have now recognized that unless the generator α and the group G are selected carefully then the exchange of information may be weak and provide almost no security.

To explain the potential problem, consider the cryptosystem described above using the group Z_p^*. The modulus p is public information that defines the cryptosystem and can be expressed as t.Q+1 with t≥2 and t relatively small. This is always possible since p is odd for large primes (i.e. t could be 2).

Let S be a subgroup of Z_p^* of order t (i.e. it has t elements, each of which is element of Z_p^*) and let γ be a base for S, i.e. each element of S can be expressed as an integral power of γ and raising γ to an integral power produces an element that is itself in the subgroup S. If α is a generator for Z_p^*, then we can take $\gamma = \alpha^Q$ without loss of generality.

If E is an active adversary in the key exchange protocol between two parties A and B then the attack proceeds as follows:

1. E intercepts the message α^x sent by A and replaces it by $(\alpha^x)^Q = \gamma^x$ and sends it on to entity B.
2. E intercepts the message α^y sent by B and replaces it by $(\alpha^y)^Q = \gamma^y$ and sends it on to entity B.
3. A computes $(\gamma^y)^x = \gamma^{xy}$.
4. B computes $(\gamma^x)^y = \gamma^{xy}$.
5. Although E does not know the key γ^{xy}, E knows that the common key γ^{xy} lies in the subgroup S of order t as γ is a generator of S. By definition γ^{xy} must produce an element in the subgroup S. Since S is of order t it has precisely t elements. If t is small enough then E can exhaustively check all possibilities and deduce the key.

Since E selects Q, t can always be taken to be 2 and so the threat is practical.

A similar attack may be mounted with cryptosystems using groups other than Z_p^* which will be vulnerable if the element selected as a base or generator generates a subgroup which itself has a small subgroup of order t.

It is therefore an object of the present invention to provide a method for checking if modification of messages has occurred or in the alternative some method to prevent the attack from being mounted.

In general terms, the present invention is based upon utilization of predefined characteristics of the order of the subgroup.

In one aspect, the base of the cryptosystem is chosen to be a generator of a subgroup of a relatively large

prime order. Substitution of any other non-unit generator is of no advantage to an attacker since it does not produce an element in a smaller subgroup that can be exhaustively searched.

In another aspect, factors of the order of the group generated by the base are used to ensure that the key does not lie in or has not been modified to lie in a proper subgroup of relatively small order, i.e. one that may feasibly be exhaustively searched by an interloper.

Embodiments of the invention will now be described by way of example only with reference to the accompanying drawings, in which

Figure 1 is a schematic representation of a data communication system.

Referring therefore to Figure 1, a pair of correspondents, 10,12, denoted as correspondent A and correspondent B, exchange information over a communication channel 14. A cryptographic unit 16,18, is interposed between each of the correspondents 10,12 and the channel 14. A key 20 is associated with each of the cryptographic units 16,18 to convert plain text carried between each unit 16,18 and its respective correspondent 10,12 into ciphertext carried on the channel 14.

In operation, a message generated by correspondent A, 10, is encrypted by the unit 16 with the key 20 and transmitted as ciphertext over channel 14 to the unit 18.

The key 20 operates upon the ciphertext in the unit 18 to generate a plaintext message for the correspondent B, 12. Provided the keys 20 correspond, the message received by the correspondent 12 will be that sent by the correspondent 10.

In order for the system shown in Figure 1 to operate it is necessary for the keys 20 to be identical and therefore a key agreement protocol is established that allows the transfer of information in a public manner to establish the identical keys. A number of protocols are available for such key generation and most are variants of the Diffie-Hellman key exchange. Their purpose is for parties A and B to establish a secret session key, K.

The system parameters for these protocols are a multiplicative group G and a generator α in the group G. Both G and α are known. Correspondent A has private key x and public key $p_A = \alpha^x$. Correspondent B has private key y and public key $p_B = \alpha^y$. Correspondent A and B exchange respective public keys and exponentiate with their private keys to obtain a common session key α^{xy}.

As noted above, the key exchange and therefore the ciphertext, is vulnerable if interloper E intercepts the transmission of α^x and α^y and raises each to the power Q.

In a first embodiment, the attack is foiled by defining the system parameters appropriately so that no advantage is provided to the interloper by performing a substitution. Moreover, the base or generator of the cryptosystem is selected so that tampering with the key exchange between A and B can be detected.

By way of example, for a public key system using the group Z_p^*, initially a subgroup S of Z_p^* is selected which has a prime order. The subgroup S of prime order q will only have subgroups of order 1 or the prime q itself. For example, if p is chosen as 139 then

$$Z_{139}^*$$

contains subgroups of order 1,2,3,6,23,46,69 and 138. Of these, the subgroups of order 2,3 and 23 are of prime order.

Accordingly, if the base used in the public key system is chosen to be a generator γ of a subgroup S of Z_p^* of prime order q rather than a generator x of Z_p^* itself, an attempt by the interloper to substitute a smaller subgroup may be readily detected.

For example, 34 is a generator of the subgroup of order 23 in

$$Z_{139}^* .$$

Therefore the base is chosen to be 34 for key exchange and generation.

The selection of the subgroup S of prime order q restricts the interloper E to an exponent of either 1 or the prime q, i.e. 23 in the example given. If the exponent is chosen to be the order q of the subgroup S then the message produced from the generator of the subgroup exponentiated to q will be the identity element, i.e. 1 in the example given. Therefore one or both correspondents may check the message and if it corresponds to the identity element it is rejected.

Selection by the interloper E of the exponent to be 1 will of course not be of use as the discrete log problem will still be intractable and provided the order of the subgroup is sufficiently large a brute force approach is impractical.

It will of course be understood that the example given of p = 139 is for illustrative purposes only and that in practical implementations the prime p will be of the order of 10^{150} and the order of the subgroup will typically exceed 10^{40}.

In a second embodiment, the order of the subgroup need not be prime and the attack is foiled by monitoring the received message. The order of the subgroup may therefore have a number of small divisors, t_1, t_2 which are sufficiently small to render the exchange vulnerable. To foil such a substitution, at least one of the correspondents A,B takes the message received from the other correspondent, i.e. α^x for B or α^y for A and raises the message to the power t for each small divisor of (p-1). If the result is 1 it indicates that a new value of the message may have been substituted, as $(\alpha^x)^{Q.t.}$ mod (p-1) will always be 1. The fact that the result is 1 is not determinative that a substitution has been made but the

probability that $(\alpha^x)^t = 1$ for large values of p is small. The key exchange can be terminated if the result is 1 and a new key exchange initiated. If with different values of private keys x and y successive key exchanges yield a result of 1 when tested above, then it is assumed that an interloper is actively monitoring the data exchange and further communication is terminated.

The determination of the value α^{xt} may be made by exponentiation of the message α^x with the possible values of t by an exhaustive search. Alternatively, given the order of the subgroup, values of the message that yield the group identity can be tabulated and a simple comparison made to determine if the message is vulnerable.

As a third embodiment, the value of p is selected to be of the form 2q+1 where q is itself a prime. The only subgroups of Z_p^* have orders 1,2,q and 2q. The generator of the subgroup of order q is selected for the key exchange so that 't' can only be 1 or q. If the subgroup of order 1 is selected then the message $(\alpha^x)^q$ will be the identity element, e.g. 1, and this can readily be checked. q will be selected to be relatively large to render an attack on the discreet log problem unfeasible.

The above techniques provide a clear indication of an attempt by an interloper to substitute a subgroup and a foil that is readily implemented by a careful selection of the generator and a check for the identity element.

The above examples have utilized the group Z_p^* but other groups may be used as noted above, for example, an elliptic curve group over a finite field. In the case of an elliptic curve over the field F_p elements where p is a prime power, there is an elliptic curve group G for each integral order lying between $p+1-2\sqrt{p}$ and $p+1+2\sqrt{p}$. With high probability, there is a prime q lying in this interval and by selecting this elliptic curve group, G_q, of order q for use in the cryptosystem, the group G_q will only have subgroups of order 1 and the prime q itself. Accordingly, selection of the group G_q will avoid substitution of subgroups of relatively small order and any attempt at substitution will not yield any benefits to the interloper.

A particularly convenient finite field is the field F_{2^m} which may be used for the generation of elliptic curve groups.

As an alternative approach to the selection of a group of prime order, the order of the elliptic curve may be chosen of order n, where n is not a prime and messages are monitored by at least one of the correspondents. The integrity of the message is verified by raising the message to the power d for each small divisor d of the order n. In this case, if the result is the group identity, typically O, then it is assumed that a substitution has been made and the transmission is terminated.

Again, therefore, a group is selected that is either of prime order to inhibit substitution or a group is chosen to have an order with small divisors. In each case, substitution can be checked by monitoring the message by at least one of the correspondents.

Similar considerations will apply in other groups and careful selection of the order of the groups utilized will provide the benefits described above.

An alternative attack that may be utilized is for the interloper E to substitute a new message "e" for that transmitted from A to B and vice versa.

The new message e is chosen to be an element of a subgroup S of the group G of low order, i.e. a relatively small number of elements. When B receives the message e he exponentiates it with his secret key y to generate the session key. Similarly, when A receives the message e he exponentiates it with the secret key x to generate the session key.

Exponentiation of an element of a subgroup will produce an element within that group so that the session keys generated by A and B lie in the subgroup S. If S is of relatively low order, there is a reasonable chance that the keys generated by A and B will be identical. In that case a message encrypted with the session key may be intercepted and the small number of possibilities that exist for the key can be tried by E.

If the keys are not identical then the failure will be attributed to system errors and a new attempt will be made to establish a key. This provides E with a further opportunity to substitute a different element of the subfield S in the transmission with a real probability that a correspondence will be established. Because of the relatively small number of possible elements, the possibilities may be exhausted and a correspondence made within the normal operating parameters of the system.

To overcome this possibility, the order of the group is selected to have factors that are either large primes or provide trivial solutions that disclose themselves upon simple examination. In the case of the group Z_p^*, a suitable form is for the value of the modulus p to be of the form 2qq'+1 where q and q' are both large primes. The subgroups S of Z_p^* will be of order 2, q or q'. Adopting a subgroup of order 2 will provide only two possible elements which can readily be checked and, if present as the session key, the session can be terminated.

The values of q and q' will not be readily ascertained due to the difficulty of factoring the products of large primes.

Even if an exhaustive attack on the subgroup of order q or q' is viable for E, such an attack will reveal itself by a large number of repeated attempts at establishing communication. Accordingly, an upper limit may be established after which communication will be terminated. The appropriate number of attempts will be based on the factors of p-1 and the nature of the communication system.

Again, therefore, the attacks by E can be resisted by checking for values of the session key that are indicative of the vulnerability of the session and by appropriate selection of the order of the group. It will be recognised that selection of the modulus of the form 2q+1 as exemplified in the third embodiment above provides the requisite robustnesss for resisting a substitution attack by E.

These techniques are also effective to prevent interloper E from taking a known public key α^a, raising it to an appropriate power such that α^{aQ} is in a small subgroup. The interloper can then determine aQ, and use this as his private key. There are situations where the interloper can use this to impersonate correspondent A and also convince a certifying authority to certify the public key α^{aQ} since the interloper E can prove he knows aQ.

In the above examples, the checking for elements lying in subgroups of relatively small order has been performed by exponentiating the message to the power of the small divisors of the order of the group. An alternative method which will indicate whether or not the message lies in a proper subgroup, without necessarily identifying the order of the subgroup, is to exponentiate the message to the order $^n/_p$ where n is the order of the group G and p ranges over all prima divisors of n. If the result is the group identity (1 in the case of Z_p^*) then it indicates that the message does lie in a subgroup. Depending upon the strategy used to determine the order of the group G, it is possible either to reject the message or to test further to determine the order of the subgroup.

Claims

1. A method of determining the integrity of a message exchanged between a pair of correspondents, said message being secured by embodying said message in a function of α^x where α is an element of a finite group S of order q, said method comprising the steps of at least one of the correspondents receiving public information α^x where x is an integer selected by another of said correspondents, determining whether said public information α^x when exponentiated to a value t where t is a divisor of q provides a resultant value α^{xt} corresponding to the group identity and rejecting messages utilizing said public information if said resultant value corresponds to said group identity.

2. A method according to claim 1 wherein a plurality of values of t are utilized and each resultant value compared to the group identity.

3. A method according to claim 1 wherein said order q is a prime number.

4. A method according to claim 1 wherein said group S is a subgroup of a group G of order n.

5. A method according to claim 4 wherein q is a prime number.

6. A method according to claim 1 wherein said group is a multiplicative group Z_p^* of integers mod p where p is a prime.

7. A method according to claim 1 wherein said group is a multiplicative group of a finite field.

8. A method according to claim 1 wherein said group is an elliptical curve group G over a finite field.

9. A method according to claim 1 wherein said group is over a finite field F_{2^m}.

10. A method according to claim 9 wherein said group is an elliptic curve group.

11. A method according to claim 6 wherein said modulus p is of the form 2q+1 and q is a prime.

12. A method according to claim 6 wherein said modulus p is of the form nqq'+1 and q and q' are relatively large primes.

13. A method according to claim 1 wherein said message is a component of a session key α^{xy} where y is an integer selected by said one correspondent.

14. A method of establishing a session key of the form α^{xy} for encryption of data between a pair of correspondents having respective private keys x and y comprising the steps of selecting a finite field Fp, establishing a subgroup S of the field Fp having an order q, determining an element α of the subgroup S to generate a relatively large number of the q elements of the subgroup S and utilising said element α to generate a session key at each corespondent of the form α^{xy} where x is an integer selected by one of said correspondents and y is an integer selected by another of said correspondents.

15. A method according to claim 6 wherein said order q of said subgroup S is a prime.

16. A method according to claim 14 including the step of receiving at one of said correspondents a message α^x, exponentiating said message α^x to a value t where t is a divisor of the order of the subgroup, comparing a resultant value α^{xt} to the group identity and preventing establishment of said session key if said value corresponds to the group identity.

17. A method according to claim 14 wherein said order of said subgroup is of the form nqq'+1 where n, q and q' are each integers.

18. A method according to claim 17 wherein the values q and q' are each prime numbers.

19. A method according to claim 18 wherein n has a value of 2.

20. A method according to claim 14 wherein said subgroup is selected to have an order that is to be a

function of the product of a pair of primes q,q' and said element α is a generator of a subgroup of an order of one of said primes q,q'.

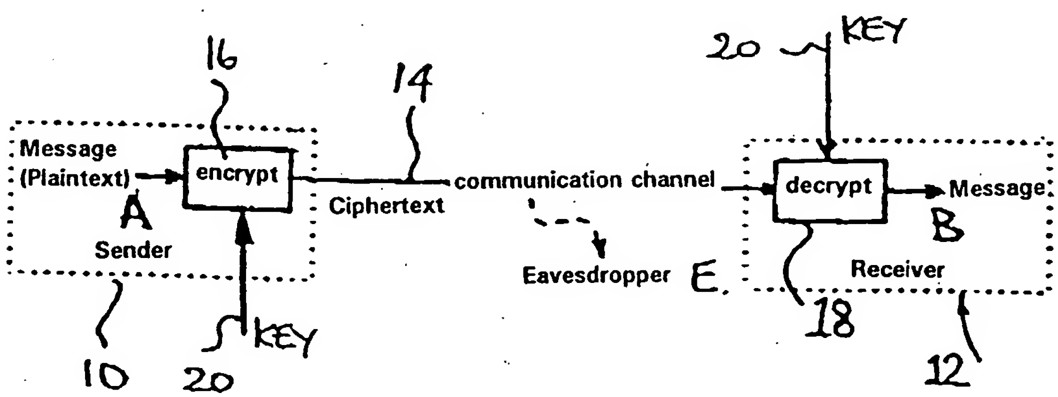
16
14
20
KEY
Message
(Plaintext)
encrypt
A
Sender
Ciphertext
communication channel
Eavesdropper
E
decrypt
Message
B
Receiver
10
20
KEY
18
12

FIG. 1